SAND2022-5364J

WILEY-VCH

Physical Compact Model for Three-Terminal SONOS Synaptic Circuit Element

Su-in Yi[1], A. Alec Talin,[2] Matthew J. Marinella[3], and R. Stanley Williams[1,*]

[1]Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA

[2]Sandia National Laboratories, Livermore, CA 94550, USA

[3]School of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ 85281, USA

* Corresponding author: rstanleywilliams@tamu.edu

Keywords: Memristor, Neuromorphic, SONOS, Compact model, Three-terminal synaptic device

[Abstract]

A well-posed physics-based compact model for a three-terminal Silicon-Oxide-Nitride-Oxide-Silicon (SONOS) synaptic circuit element is presented for use by neuromorphic circuit/system engineers. Based on Technology Computer Aided Design (TCAD) simulations of a SONOS device, the model contains a non-volatile memristor with the state variable Q_M representing the memristor charge under the gate of the three-terminal element. By incorporating the exponential dependence of the memristance on Q_M and the applied bias V for the gate, the compact model agrees quantitatively with the results from TCAD simulations as well as experimental measurements for the drain current. The compact model was implemented through VerilogA in the circuit simulation package Cadence Spectre, and reproduced the experimental training behavior for the source-drain conductance of a SONOS device after applying writing pulses ranging from -12V to +11V, with an accuracy higher than 90%.

Improvements in computing performance and efficiency recently opened up opportunities[1] in artificial intelligence (AI), represented by the seminal work of Alex Krizhevsky et al.[2] in 2012 where graphic processing units (GPUs) combined with big data excelled on an image classification task (Top-5 error rate of 15.3 % achieved by artificial neural networks versus

26.2 % from the runner-up algorithms). This achievement reignited the attention for deep neural networks dating back to the 1980's[3, 4], in particular for convolutional neural networks (CNN). Further efforts to optimize computing hardware are being actively pursued, such as using field programmable gate arrays (FPGAs) and application-specific integrated circuits (ASICs)[5, 6], including the tensor processing units (TPUs)[7, 8] developed by Google and system on a chip (SoC) scale AI-accelerators used in mobile phones. This recent progress in computing hardware is still bound to CMOS technology, where the von Neumann bottleneck[9] limits speed and power efficiency. In-memory analog computing based on Ohm's and Kirchhoff's Laws[10-12], utilizing non-volatile-memories (NVMs), such as resistive random access memory (RRAM)[13-15], phase change memory (PCM)[16, 17] and Flash memory,[18-21] is a promising route to eliminate the von Neumann bottleneck and implement neuromorphic computing circuits for future AI systems[22-26].

Here we present a dynamical compact model for three-terminal SONOS (Silicon-Oxide-Nitride-Oxide-Silicon) synaptic circuit elements. Three-terminal memory devices such as floating gate (FG) or SONOS Flash memories have a long history for manufacturable data storage applications, but their compact models have relied on static MOSFET behavior[27-29]. Moreover, a significant number of three-terminal synaptic devices are now being reported with various material systems[30-32], but most publications describe experimental characterization without supplying a compact model[33], so there exists a significant gap between the device community and circuit designers. Our work addresses this gap by constructing a well-posed compact model for three-terminal synaptic circuit elements[34-36], beginning with the technologically mature SONOS device, which is a staple for NAND Flash memory products[37-40]. In this work, we utilized Technology Computer Aided Design (TCAD) physics-based calculations within Synopsys Sentaurus to simulate the physics of a device, and identified a key state variable Q_M, the amount of charge in the SONOS trap-layer, to guide us in constructing a compact model. Subsequently, we validated the model through a circuit simulation using Cadence Spectre to compare with the experimentally measured behavior of source-drain current after applying a wide range of voltage pulse amplitudes on the gate of a SONOS device. The use of both physics-based simulations and experimental data to identify the state variables and calibrate the model was crucial, since the critical state variable Q_M is extremely difficult to measure experimentally (see the Supplementary Video).

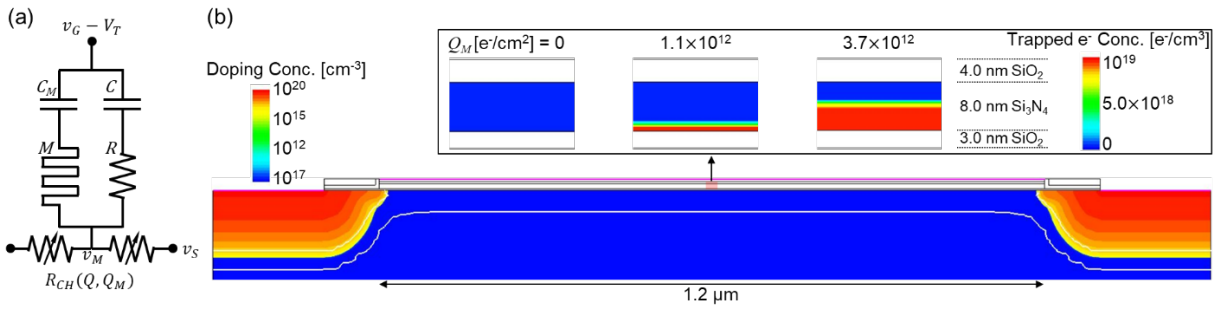

Figure 1. (a) The equivalent circuit diagram for a three-terminal SONOS synaptic circuit element based on physical state variables Q_M and Q. The state variable Q_M is the total amount of charge that has passed through an effective memristor M, equal to the charge on the series capacitor C_M (the branch on the left), and controls the slow dynamics. The memristance is a function of both the state variable Q_M and the applied voltage, v_G -V_T -v_M, with a strong nonlinearity. The parallel branch on the right, composed of a series resistor (R) and capacitor (C), is responsible for the short-term dynamics with a state variable Q, which represents the typical switching behavior of a MOSFET. For simplicity, the capacitor C is described as a linear capacitor, which holds when v_G -V_T>0, where the influence of charge depletion is negligible. R describes an effective resistance including the contact resistance of the gate electrode, the scattering by the ionized impurities in the depletion region, etc. The conductances and currents through the channel of the SONOS device are represented by variable resistors with identical resistances R_{CH}, and depend on the two state variables Q and Q_M. (b) The geometry of a SONOS device in TCAD simulations and experiments (Ref.[33]). Color plot shows the doping concentration of 2-D cross section perpendicular to the gate width. Inset shows three exemplary programmed states with color plots of trapped electron concentration in Si_3N_4 layer.

Figure 1 (a) depicts the equivalent circuit diagram for the well-posed compact model of a three-terminal SONOS synaptic circuit element[34-36] for neuromorphic applications, which consists of three components. The first is found on the left vertical branch, composed of a series capacitor C_M and an extended memristor[41] M, defined as

$$M = \mathrm{f}(Q_M, V) \text{ and} \tag{1a}$$

$$\frac{dQ_M}{dt} = I_M = g(Q_M, V). \tag{1b}$$

where we are not directly interested in the current I_M through the memristor, which is in general too small to be measured experimentally, but rather how the memristor charge controls the long-term SONOS dynamics for synaptic weight modulation or non-volatile memory through the time dependence

$$Q_M = C_M(v_G - V_T - v_M)\left(1 - e^{-t/(MC_M)}\right), \quad (2)$$

which is discussed further below. The second vertical branch on the right has a series resistor R and a capacitor C, which control the short-term dynamics represented as

$$Q = C(v_G - V_T - v_M)\left(1 - e^{-t/(RC)}\right), \quad (3)$$

and in turn the drain current, i_D is

$$i_D = \mu v_D \frac{1}{L^2} Q \quad (4)$$

Equation (4) is essentially quasi-static[42, 43], without an explicit time dependence

$$i_D = \mu v_D \frac{1}{L^2} C(v_G - V_T - v_M) = \mu v_D \frac{W}{L} C_{ox}(v_G - V_T - v_M) \quad (5)$$

due to the small time constant $\tau = RC \ll 1$ns. We performed circuit simulations based on a SONOS device with W=1.2 μm, L=7 μm, μ=350 cm^{2}V^{-1}s^{-1}, and C=26 fF[18, 33] as shown in Fig.1 (b) combined with R=10^8 Ω, chosen to manage the simulation speed with a time constant τ=2.6 μs, which permits the use of a timestep as large as 1ns in Cadence Spectre to examine both short and long term dynamics. For most practical circuit simulations, the short-term dynamics can be safely eliminated by making Q a parameter instead of a state variable[18], as discussed in Section 2 of the Supplementary Information. The last component comprises two identical variable resistors, derived from v_D/i_D as

$$R_{CH} = \left[\mu \frac{1}{2L^2}(Q - \chi Q_M)\right]^{-1}, \quad (6)$$

where χ stands for the percentage of Q_M contributed by the silicon channel. The memristor conductance is a function of the Q_M history as described below.

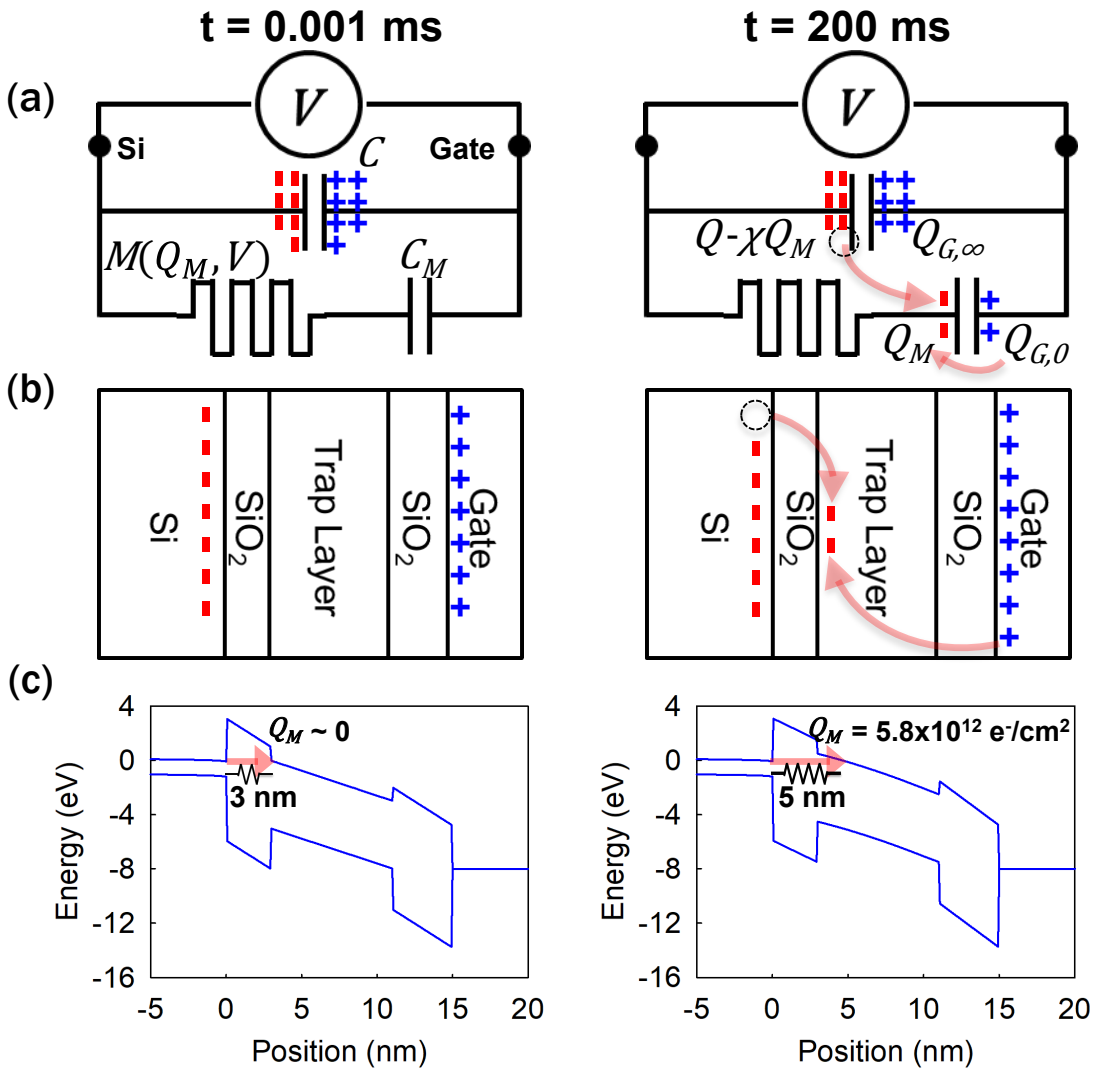

Figure 2. Synaptic weight dynamics modeling of a SONOS device with a memristor and two capacitors. The amounts of charge on the capacitors and the corresponding band diagrams are taken from TCAD simulations with a bias of V= 8 V as an example. Two distinct points of time at t= 1 μs (left) and t= 200 ms (right) after the bias is applied are shown. Each '-' symbol represents roughly 2.5×10^{12} electrons per cm^2 (e^-/cm^2) (a) Equivalent circuit diagrams of a SONOS layer stack with two terminals: the gate and silicon channel (source and drain are grounded). A memristor M and a capacitor C_M account for the synaptic weight modulation based on the state variable Q_M. When t= 200 ms, Q_M increased so that $Q - \chi Q_M$ as well as the synaptic weight (conductance across source and drain) was reduced. (b) Schematic physical structure of the SONOS stack. The thicknesses of the ONO layers are 3 nm, 8 nm, and 4 nm, respectively, whereas the silicon channel and gate are much larger. (c) Corresponding band diagram across the SONOS layer at t= 1 μs compared to t= 200 ms, which exhibits a thicker effective tunneling barrier (5 nm) that is larger than the tunneling oxide (3 nm). The memristor M with the state variable Q_M equivalently reflects the varying tunneling barrier thickness.

We performed TCAD simulations to determine the quantitative dynamics of the state variables Q_M and Q, representing charge transport via tunneling and electron drift-diffusion[44], because they are not easily experimentally accessible. A memristor circuit element was necessary to describe the dynamics of Q_M, whereas a linear *R*-*C* circuit was suitable for Q. The results in Fig. 2 provided the information used to derive the compact model in Fig. 1. To focus on the dynamics of synaptic weight modulation, the drain and the source were grounded and denoted as 'Si', which eliminated the channel current. A bias of V= 8V was applied to the gate at t= 0, with Fig. 2 representing two times after bias application, t= 1 μs and t= 200 ms. The resistor in the short-term branch (R) in Fig. 1 was not included because C was fully charged within one microsecond in the TCAD simulations. The amount of charge under 8 V bias at t= 1 μs (left, Fig. 1 (a)) was Q= 17.5×10^{12} e$^-$/cm^2, corresponding to 2.8 μC/cm^2, which shows good agreement with the expected value of 2.5 μC/cm^2 calculated from 8 V on a capacitor with 0.31 μF/cm^2 (3 nm SiO_2, 8 nm Si_3N_4, and 4 nm SiO_2) visualized schematically in Fig. 2 (b). The area density of electrons was used as the unit for Q and Q_M for more intuitive analysis in this paper. A slightly larger amount of charge compared to the expected capacitor value is attributed to the negative threshold voltage V_T = -0.01V and nonlinear Q-V curve near the depletion region ($v_G \cong V_T$) of a MOSFET. The corresponding band diagram across the material stack is shown in Fig. 2 (c). The trap layer (Si_3N_4) remains electrostatically neutral until t= 1 μs with Q_M= 0, and provides a tunneling barrier thickness of 3 nm due to the 9 eV forbidden gap of SiO_2. After allowing enough time for tunneling, at t= 200 ms (Fig. 2 right column), Q_M= 5.8×10^{12} e$^-$/cm^2 and Q_G= 20×10^{12} h$^+$/cm^2, hence Q - χQ_M= 14.2×10^{12} e$^-$/cm^2 because of charge neutrality (Q_G= Q_M + Q - χQ_M). Consequently, the synaptic weight of the SONOS decreased based on the net charge changing from 17.5×10^{12} e$^-$/cm^2 at t= 1 μs to 14.2×10^{12} e$^-$/cm^2 at t= 200 ms. In other words, 57% of the increased Q_M (0→5.8×10^{12} e$^-$/cm^2) is contributed by the electron charge on C (= Q - χQ_M : 17.5×10^{12} e$^-$/cm^2→14.2×10^{12} e$^-$/cm^2), realizing the synaptic weight depression. Although the tunneled charge is 5.8×10^{12} e$^-$/cm^2, intuitively implying Δ(Q - χQ_M)= -5.8×10^{12} e$^-$/cm^2, the loss of charge on C is spontaneously complemented by the gate electrode (Q_G: 17.5×10^{12} h$^+$/cm^2→20×10^{12} h$^+$/cm^2) due to the electrostatic force from the trapped charge located at the middle of dielectric stack (O-N-O). The Supplementary Video provides a dynamic circuit illustration of this mechanism and the relevant discussion on χ follows later (Fig. 6). The

underlying physics of the memristor in our compact model is revealed in the band diagram at t= 200 ms (right panel of Fig. 2 (c)). The slightly lifted conduction band edge (or electrostatic potential) due to the injected negative space charge (trapped electrons) alters the tunneling thickness so that an electron, from the conduction band edge of silicon interfaced with the SiO_2 tunneling oxide, encounters a thicker tunneling barrier of 5 nm compared to 3 nm at t= 1 μs when Q_M= 0. Consequently, the tunneling current decreases, i.e., $[\mathrm{d}Q_M/\mathrm{dt}]_{Q_M=5.8\times10^{12}} < [\mathrm{d}Q_M/\mathrm{dt}]_{Q_M=0}$. Figure 2 presents results at two discrete times, but the memristance changed smoothly and continuously with the state variable Q_M.

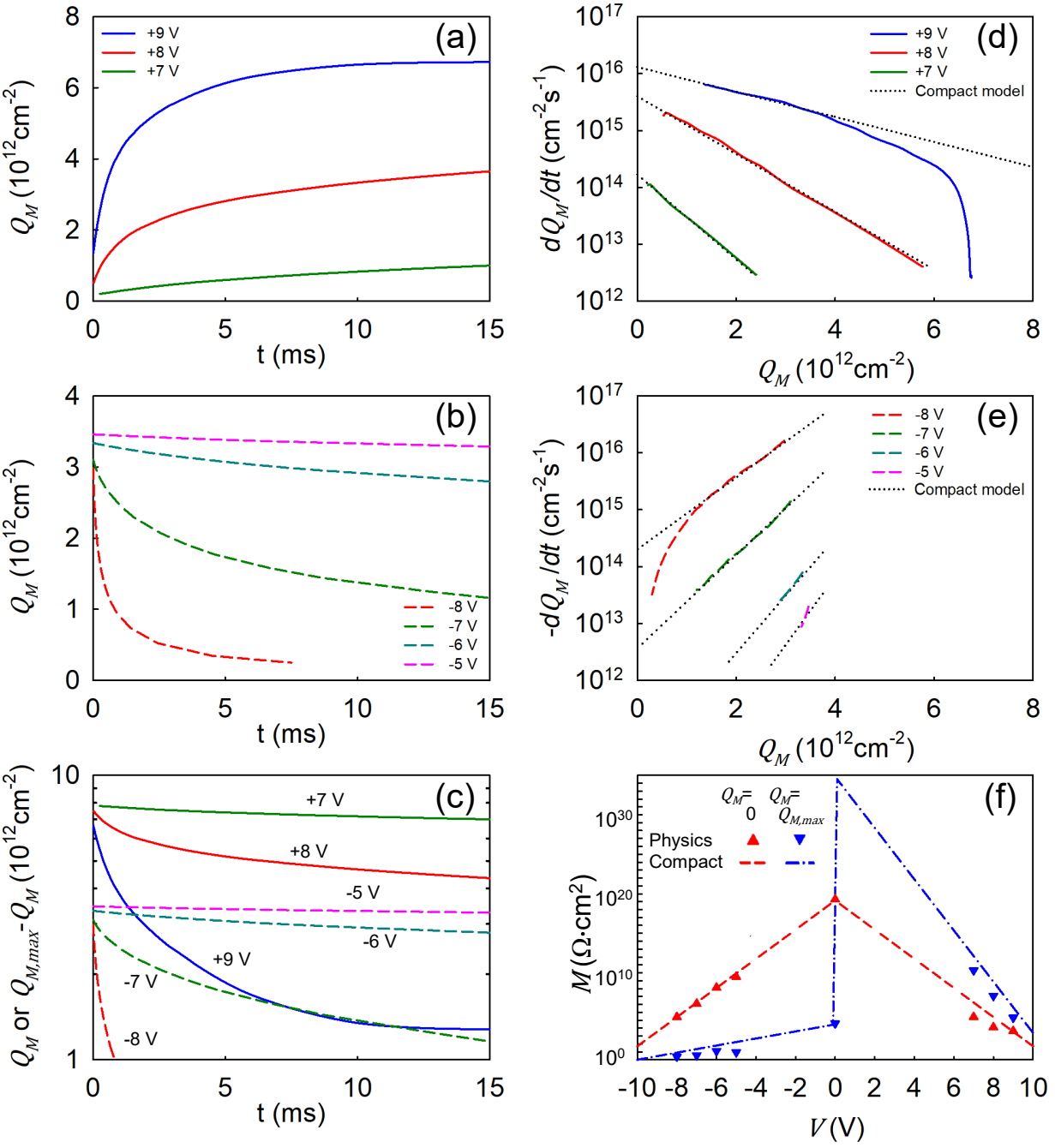

Figure 3. Physics model (TCAD) simulation results of a SONOS device and corresponding compact model. (a and b, positive and negative gate biases, respectively) In (a), the state variable Q_M increases with time and saturates at $Q_M = Q_{M,max} = 8\times10^{12}$ e$^-$/cm^2, which is defined by the trap concentration (10^{19} cm^{-3}) chosen for the 8 nm-thick Si_3N_4 layer. (c) Logarithmic scale plots of Q_M for $V > 0$ (or $Q_{M,max}$- Q_M for $V < 0$), conveying the identical information of (a) and (b), but emphasizing the nonlinearity of the response. The decaying slopes (=-1/τ) with respect to time show that the characteristic time is not constant. (d and e) Dynamic route maps (DRMs) extracted from (a) and (b), respectively. The nearly linear relations between log(dQ_M/dt) and Q_M demonstrate that the memristance depends exponentially on Q_M. The compact model linear approximations to the DRMs are denoted with dotted lines. The

eccentric behaviors observed for 9 V with $Q_M > 5\times10^{12}$ e$^-$/cm^2 and -8 V with $Q_M < 10^{12}$ e$^-$/cm^2 are due to Q_M approaching $Q_{M,max} = 8\times10^{12}$ e$^-$/cm^2 and $Q_{M,min} = 0$, respectively. (f) Memristance as a function of voltage. Upward and downward triangles represent the memristance values at two extremes, $Q_M = 0$ and $Q_M = 8\times10^{12}$ e$^-$/cm^2, respectively, from (d and e). The overall trend of the memristance M versus Q_M and V exhibits the dependence on the state variable (governed by a parameter γ) and a strong nonlinearity with V (governed by a parameter β) in Eq. (8).

Figures 3 (a) and (b) show the simulation results of Q_M as a function of time under constant biases: +9, +8, and +7 V in (a) and -8, -7, -6, and -5 V in (b). The dynamics are qualitatively reminiscent of an *R*-*C* circuit, where the initial current, $i_{t=0}$, is equal to V/R and Q increases inverse-exponentially until it is saturated at $Q=CV$, but is quantitatively different. For example, the two curves with +9 V and +8 V in Fig. 3 (a) show that $+9/[dQ_M/dt]_{t=0}$ from the former is not identical to $+8/[dQ_M/dt]_{t=0}$ from the latter, immediately identified by the slope of the curves at t=0. Therefore, the resistive component in series with the capacitor has a voltage dependence. Secondly, the slope of the tangent decreases exponentially (not linearly) with Q_M, which is evident from the replotted curves on a log scale in Fig. 3 (c), demonstrating that the resistance is also a function of Q_M, as forecast by Eq. (1). The y-axis represents $Q_{M,max} - Q_M$ for $V>0$ and Q_M for $V<0$, because $-t/(RC)$ is equal to $\ln(CV-Q)+D$ for $V>0$ and $\ln(Q-CV)+D$ for $V<0$, where D is a constant (see Eq. (S8)). For a usual RC circuit, $\ln(Q-CV)$ for $V<0$ (or $\ln(CV-Q)$ for $V>0$) is a linear function of time with a slope $-1/\tau$, where τ is the time constant ($= RC$, see Eq. (S5)~(S8) in the Supplementary Information for more details). However, the observed results exhibit significant nonlinearities for all seven biases, demonstrating that the resistance is a function of Q_M, which is the memristance. The dynamic route maps (DRMs) in Figs. 3 (d) and (e) describe the dependence of the dynamics on the state variable. The Q_M dynamics results obtained from our physics simulations were fitted by a *M*-*C* circuit as shown by the dotted lines in Figs. 3 (d) and (e), with the dynamical equation:

$$\frac{dQ_M}{dt} = -\frac{1}{M(Q_M,V)C_M}Q_M + \frac{V}{M(Q_M,V)} \quad (7)$$

This equation breaks down when Q_M approaches the extrema at 0 or $Q_{M,max}$, which both require extreme times to reach. As long as the SONOS operation range is limited to intermediate states with moderate values of Q_M, the compact model featuring the exponential correlation between M and Q_M can emulate the behavior with only small errors. The compact model calibrated to an experimental SONOS device (Fig. 8) showed that Q_M was bounded between 20 fC to 60 fC even

after many pulses of various magnitudes, because dQ_M/dt changes exponentially with Q_M. The fitted memristance values at Q_M=0 and Q_M= $Q_{M,max}$=8×10^{12} e$^-$/cm^2 are marked as upward and downward triangles, respectively in Fig. 3 (f) for seven different bias points performed in our physics simulation. Seven sets of memristance values as a function of Q_M were interpolated and extrapolated to establish the following analytic equation:

$$M(Q_M, V) = \alpha \times \beta^{10-|V|}\left(\gamma \times \delta^{10-|V|}\right)^{\frac{(-1)^n Q_M}{8\times 10^{12}}} \quad (8)$$

where α= 50, β= 70, γ= 50, δ= 25, and n is 0 for V>0 and 1 for V<0. The dependences on V and Q_M are described by β and γ, respectively. Additionally, it can be found that the sensitivity of the memristance on Q_M (slopes of the dotted lines in Figs. 3 (d) and (e)) also exponentially increases with V, which is reflected by the fitting parameter δ. Although the curves from the analytic equation denoted by lines produce some errors compared to the physics simulation results (symbols), it can account for all possible combinations of Q_M and V so that it can stand alone as a circuit element. A discontinuity in M between V= 0- and V= 0+ is observed when Q_M= $Q_{M,max}$, for which the physical mechanism is attributed to a change in the charge tunneling dynamics. While a larger Q_M provides a thicker tunneling barrier in the case of trapping into Si_3N_4 (V>0), it fosters detrapping when V<0 because the electrostatic potential of the trap is higher compared to when Q_M= 0.

Figures 4 (a) and (b) reveal the familiar pinched hysteresis loops characteristic of memristors with an applied sinusoidal voltage. In the present system, both loops are traversed clock-wise, which is different from most familiar memristors that show both clock-wise and counter clock-wise trajectories depending on the sign of the voltage. Although the I_M-V characteristic of in Fig. 4 (b) appears to be that for a volatile memristor, the expanded view in Fig. 4 (c) shows that there are two different slopes at the zero crossing, showing that the memristor is nonvolatile but that the states near the zero crossing have a very large resistance compared to those at high voltage amplitudes. Figure 4 (d) shows that there is an avoided zero crossing when there is a capacitor in series with the memristor, which would be the case for an experimental measurement of this system since the capacitance is intrinsic to the structure. Blue solid curves in Figs. 4 (a) and (b) were obtained by TCAD simulations to verify the feasibility of our compact model. The compact model and physics simulation present good agreement, although they have a quantitative

mismatch owing to the simple analytical form of Eq. (8) that can be improved at the cost of complexity. Red solid curves in Figs. 4 (a) and (b) represent the long-term gate current, $I_{G,0}$, that is the sole way to deduce I_M through experimental measurements. The shape of $I_{G,0}$ is almost identical to I_M, but with different magnitudes, approximately half of I_M, which is consistent with the illustration in the Supplementary Video and justifies the parameter χ ranging from 0 to 1. The detailed procedure for extracting the long-term gate current ($I_{G,0}$) from the total gate current ($I_{G,0}$ + $I_{G,\infty}$) is available in Figure S2. Memristors' nonideal writing characteristic, so called nonlinearity and asymmetry, which is reflected by 'fading memory effect'[45], was also found from simulations of our compact model under sinusoidal voltage biases (Figure S3).

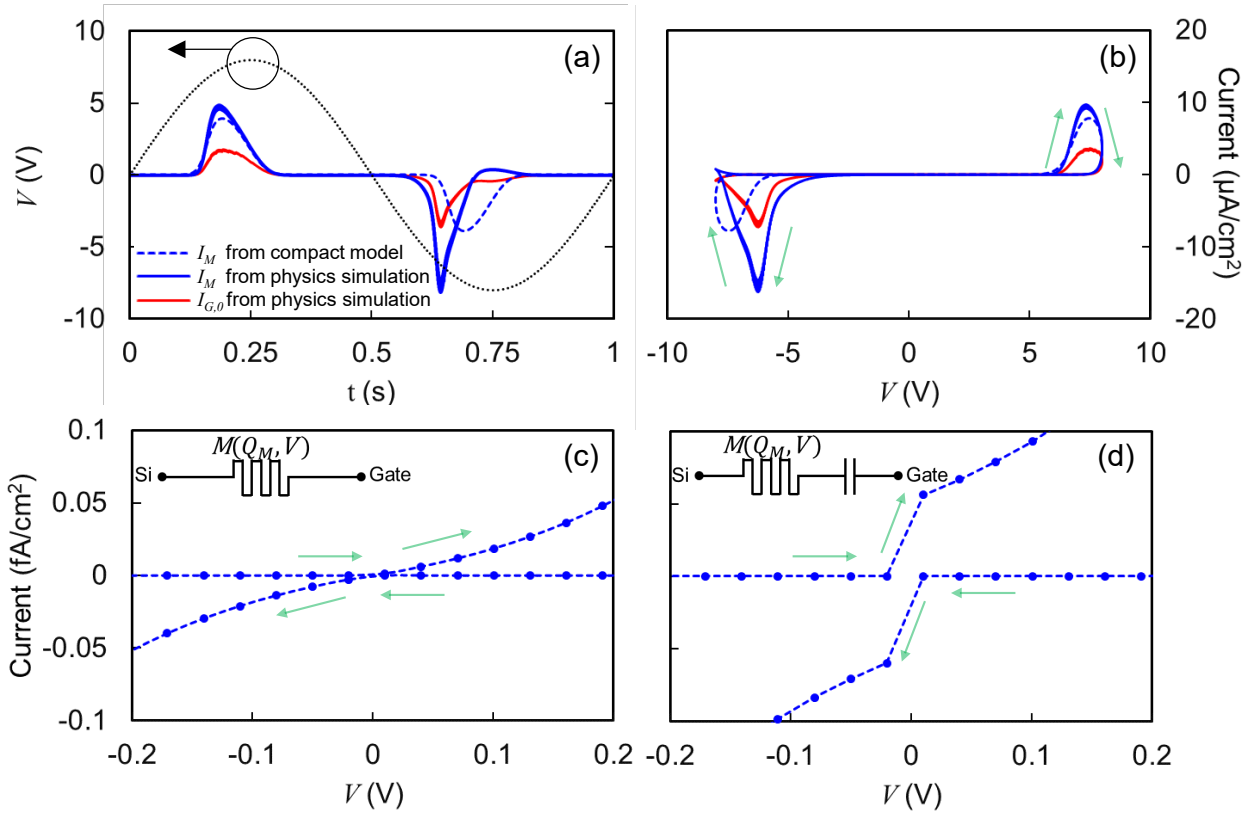

Figure 4. The characteristic of memristor in the branch of long-term dynamics under an AC voltage. (a) Current (blue: I_M and red: $I_{G,0}$) as a result of an AC voltage of 8 V and 1 Hz. Dashed-line is I_M from the compact model of Eq. (7) and (8), whereas solid lines are from physics simulation of Synopsys Sentaurus. (b) Current-voltage hysteresis curve of the memristor, I_M, and the experimentally observable long-term gate current, $I_{G,0}$, replotted from (a). The trajectories of hysteresis curve are always clockwise, whereas general oxide memristors exhibit bipolar trajectories (e.g., clockwise with V>0 and counter-clockwise with V<0). (c) Hysteresis curve of the memristor near

the zero crossing. (d) Hysteresis curve of a series connection of a memristor and a capacitor near zero crossing, which is a magnified view of (b).

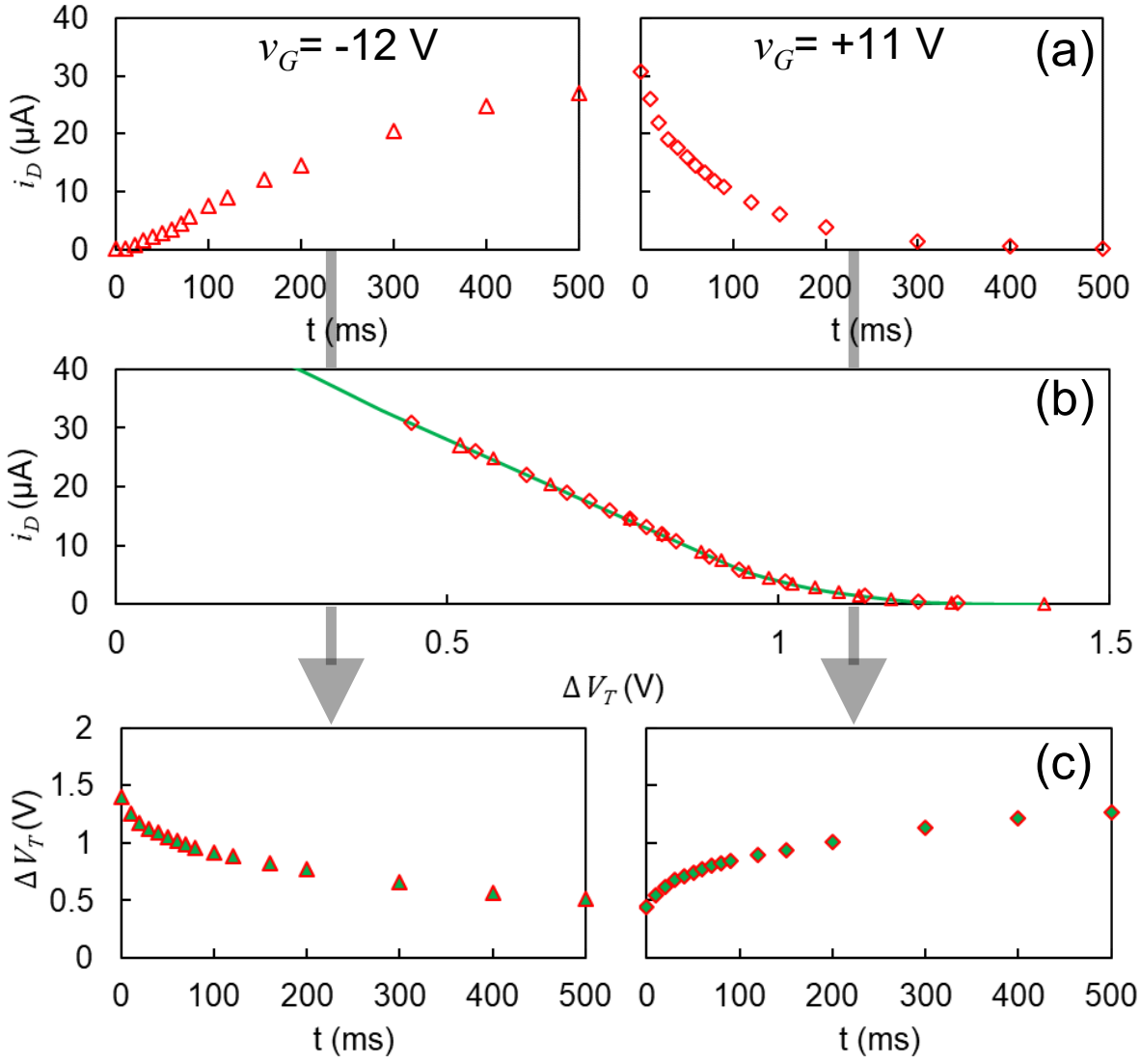

Figure 5. Facile calibration protocol applicable to general three-terminal synaptic circuit elements by utilizing the threshold voltage V_T of a MOSFET. The experimentally measured i_D after every pulse of v_G=-12 V and +11 V of the SONOS device from Ref. [33] is employed to illustrate the procedure. (a) Channel current, i_D , (with v_D- v_S= 0.1 V and v_G= 2.4 V during reading) with time under repeating training pulses of v_G= -12 V (left) or v_G= +11 V (right) on the gate. (b) The one-to-one mapping of the current, i_D, versus ΔV_T is possible because the i_D -v_G curve is simply shifted on the v_G-axis, while the entire shape is negligibly deformed. (c) Based on the conversion relation in (b), ΔV_T versus time can be obtained.

We next calibrated the compact model extracted from TCAD simulations using experimental data from a SONOS device[33]. Since an accurate compact model requires dynamical information from the target, we designed a protocol to extract the state variable from measurements that can actually be performed[46], as shown in Fig. 5 for a general three-terminal synaptic device[30-32, 47, 48]. For SONOS, the measured channel currents were converted to the change in threshold voltage (V_T) so that the desired state variable Q_M could be obtained. Figure 5 (a) shows the channel current of a SONOS device potentiated by -12 V and suppressed by +11 V from the work of Agarwal et al.[33]. By utilizing the rigid shape of i_D-v_G regardless of the shift in V_T, ΔV_T can be deduced from the current based on the measured i_D-v_G curve at an arbitrary state, as shown in Fig. 5 (b), to obtain Fig. 5 (c).

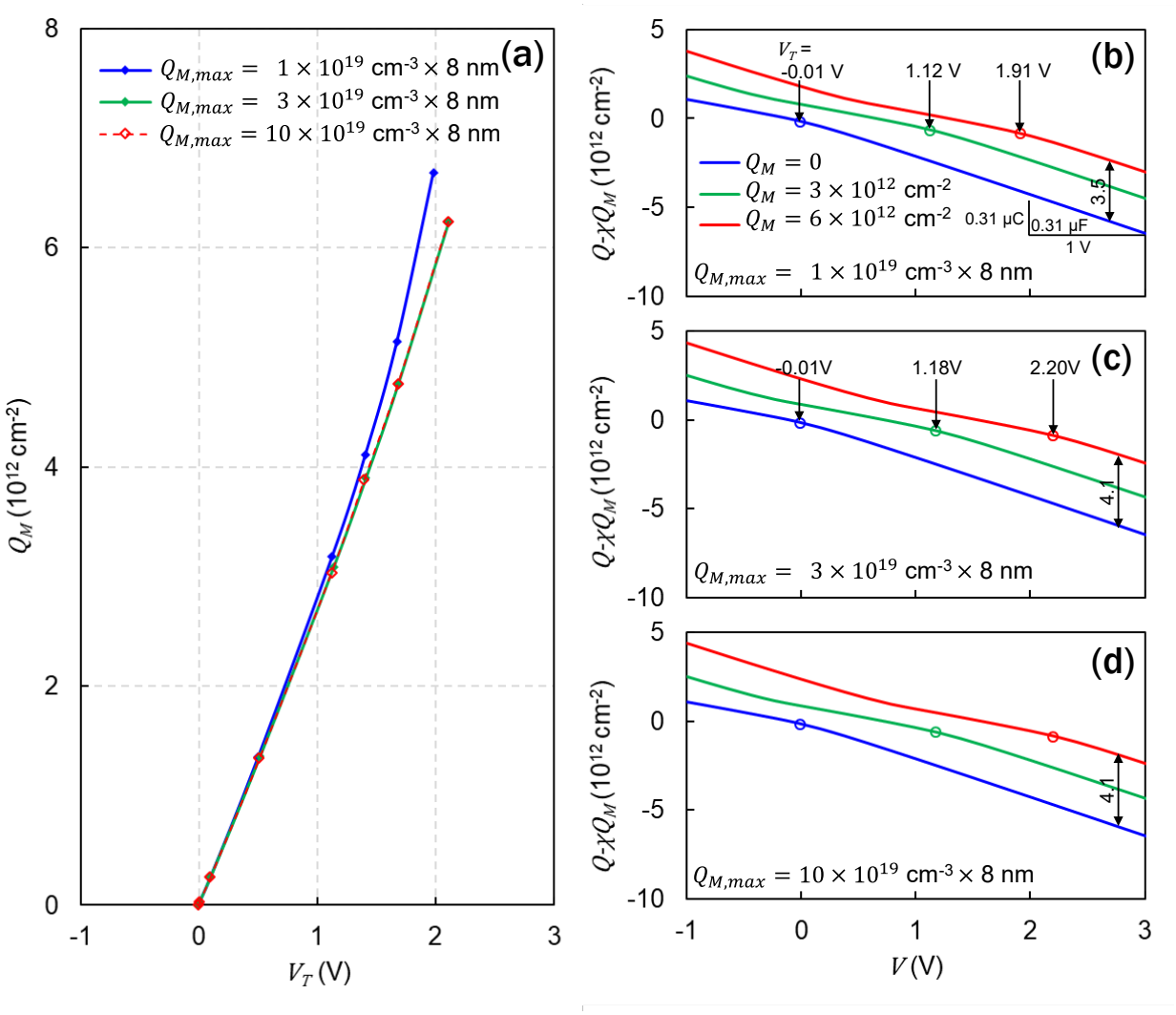

Figure 6. Relationship between V_T and Q_M under various conditions of Si_3N_4 trap density to obtain Q_M from the experimentally obtained V_T data. (a) V_T linearly increases by 1 V for every ΔQ_M= 2.5×10^{12} e$^-$/cm^2. (b) A detailed analysis for the case of trap density equal to 1×10^{19} cm^{-3}. The slope of 'Q - χQ_M' versus V relation is consistent with the expected capacitance of the ONO layer (~0.31 μF/cm^2), which governs the correlation between ΔV_T and $\chi \Delta Q_M$. For a MOSFET, ΔQ_M= 1.93×10^{12} e$^-$/cm^2 is required to induce ΔV_T= 1 V (= [0.31 μF/cm^2] / [1.6×10^{-19} C]) with χ=1 when the fixed charges exist at the interface between the silicon channel and SiO_2. For the SONOS device, for which the trapped charge is dominantly in the Si_3N_4 layer, χ= 0.73 was derived from the simulation results: ΔQ_M= 2.65×10^{12} e$^-$/cm^2 for every ΔV_T= 1 V. (c) When the trap density increases to 3×10^{19} cm^{-3}, the trapped charge has a narrower distribution, spatially closer to the channel rather than the gate. As a result, $\Delta(Q - \chi Q_M)$ contributes 77% of ΔQ_M (i.e., , χ= 0.77) and the conversion factor slightly decreased to 2.52×10^{12} e$^-$/cm^{2}V^{-1}. (d) Further increase to 10×10^{19} cm^{-3} does not create a notable change compared to 3×10^{19} cm^{-3} from (c) (i.e., χ= 0.77), hence the conversion factor remains as 2.52×10^{12} e$^-$/cm^{2}V^{-1}.

While a simple correlation between the amount of fixed charge (ΔQ_{NOT}) at the interface of the silicon channel and SiO_2 and the shift in threshold voltage (ΔV_T) holds from $\Delta V_T = \Delta Q_{NOT}/C$ for a MOSFET, when it comes to a SONOS device with ΔQ_M (most charges exist in Si_3N_4) instead of ΔQ_{NOT}, the relation becomes $\Delta V_T < \Delta Q_M/C$.[19, 49] Due to electrostatic equilibration, the trapped charge closer to the channel results in the larger ΔV_T, mainly dictated by the amount of charge at the channel. Likewise, trapped charge closer to the gate eventually draws more charge from the gate rather than from the channel (because the gate supplies the charge to the channel, a mechanism animated in the Supplementary Video), so that ΔV_T is reduced. In order to estimate χ, ranging from 0 to 1, for a generalized correlation of $\Delta V_T = \chi \Delta Q_M/C$, a series of additional TCAD simulations were performed with three different scenarios for the trap concentrations in Si_3N_4: 1×10^{19} cm^{-3}, 3×10^{19} cm^{-3} and 10×10^{19} cm^{-3}. As shown in Fig. 6 (a), all three cases require approximately $\Delta Q_M = 2.5\times10^{12}$ e^-/cm^2 to induce $\Delta V_T = 1$ V, hence the resultant χ is calculated to be 0.77, because $C = 0.31$ μF/cm^2. For the lowest trap concentration, 10^{19} cm^{-3} (Fig. 6 (b)), a slightly larger amount of trapped charge, $\Delta Q_M = 2.65\times10^{12}$ e^-/cm^2, is required to achieve $\Delta V_T = 1$ V (i.e., $\chi = 0.73$) and the incremental efficiency becomes worse for the larger Q_M's (>4×10^{12} e^-/cm^2). This is because of the limited capacity of the trap layer so that for $Q_M > 4\times10^{12}$ e^-/cm^2, for instance, the trapped charges reside not only near the interface of Si_3N_4, but also in the bulk region of Si_3N_4, which is closer to the gate and induces a smaller ΔV_T. With the higher trap concentrations in Si_3N_4 (3×10^{19} cm^{-3} and 10×10^{19} cm^{-3}) as shown in Figs. 6 (b) and (c), the newly trapped electrons always occupy the interface between Si_3N_4 and the tunneling SiO_2 layer, so that the efficiency remains the similar with $\chi = 0.77$. Based on the analysis with different trap concentration scenarios, a conversion factor of 3×10^{12} e^-/cm^2V^{-1} was chosen to extract Q_M from the experimental measured data of ΔV_T in Fig. 5 (c).

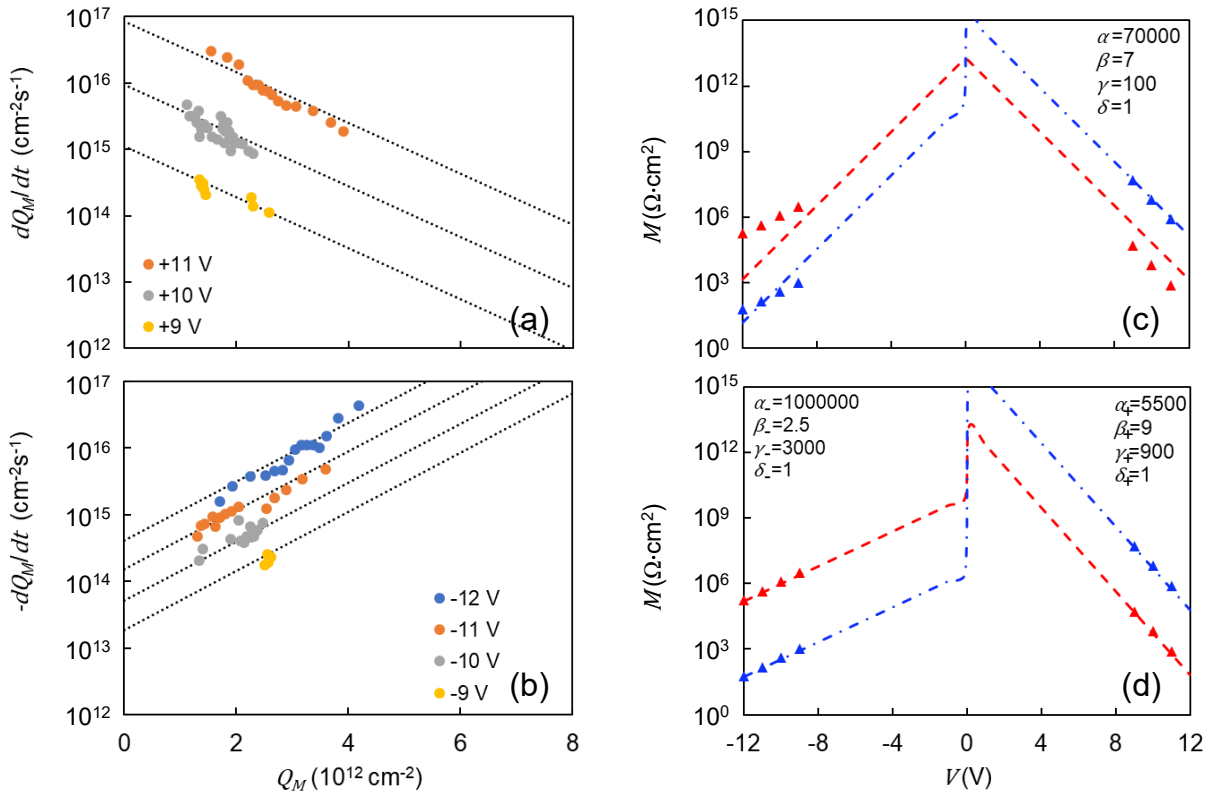

Figure 7. Extracted DRMs and corresponding memristance value maps of the experimental SONOS device in Ref. [33]. (a and b) Under both positive and negative biases for depression and potentiation, respectively, a linear approximation for log(dQ_M/dt) versus Q_M holds reasonably well, consistent with the physics simulation results and the compact model. (c) Selected experimental M values (triangles, red is Q_M=0, blue is Q_M=$Q_{M,max}$, similar to Fig. 3 (f)) to build the dotted lines in (a) and (b). The analytical function with a parameter set (α, β, γ, δ)= (70000, 7, 50, 1) simultaneously fits both potentiation and depression, but with significant errors possibly due to different physics for charge trapping and detrapping. (d) A modified function for M, where the parameters are different for potentiation (α_-, β_-, γ_-, δ_-)= (10^6, 2.5, 3000, 1) and depression (α_+, β_+, γ_+, δ_+)= (5500, 9, 900, 1), models the experimental SONOS data with negligible error.

Figures 7 (a) and (b) show the resultant DRMs from Q_M of the experimentally measured channel currents versus the number of square writing pulses on the gate[33] with seven different bias conditions: +11, +10, and +9 V for depression, and -12, -11, 10, and -9 V for potentiation. Although the experimental points are noisy compared to those from TCAD simulations, the exponentially changing memristance with Q_M and the nonlinearity with the bias voltage are present, similar to our observations from the physics simulations. The triangles (blue for Q_M= 0

and red for $Q_M = Q_{M,max}$) in Figs. 6 (c) and (d) correspond to the bias voltages in Figs. 6 (a) and (b) in a similar manner to Fig. 3 (f). The best fitting parameter set was found to be (α, β, γ, δ)= (70000, 7, 50, 1) for the experimental SONOS device compared to (α, β, γ, δ)= (50, 70, 50, 25) from the SONOS physics simulations. The significantly larger α represents a much slower tunneling by three orders of magnitude for the experimental SONOS devices at V= 10 V, implying that the tunneling mass in the TCAD simulations is too low, where the default value of Synopys Sentaurus 0.36 m_0 was employed. However, the memristance at V=0 (see Fig. 3 (f)) shows the inverse trend, such that $M = 1.41\times10^{20}$ Ωcm^2 from the TCAD model and $M = 1.98\times10^{13}$ Ωcm^2 from the experimental SONOS data, implying that the tunneling mass may need to be smaller than 0.36 m_0. This is likely due to the absence of trap-assisted-tunneling (TAT) in the TCAD simulations that becomes prominent at smaller V.[19] The nonlinearity parameter β for the experimental SONOS device is 10 times lower, so that the low bias dynamics is faster compared to the simulation model. This may also be caused by the absence of TAT through defect states in the forbidden bands of SiO_2 and Si_3N_4, and thus worse predictability for low bias cases. This is a known problem for industrial SONOS TCAD models, where the pass disturb simulations with V= 7~8 V typically underestimate Q_M compared to fabricated devices[19], because defects in SiO_2 and Si_3N_4 are difficult to model. The parameter γ, responsible for the sensitivity to Q_M, was found to be the same for both the physics simulations and the experimental data. Lastly, δ, which handles the Q_M sensitivity of M depending on V, was found to be unity, i.e., there was no noticeable dependence of Q_M on V from the experimental SONOS devices. Figure 7 (d) shows the improved fits with distinct parameters for potentiation (V<0) and depression (V>0), such that little error remains between the extracted memristance values and the analytic equation, as also found in two-terminal compact models[46].

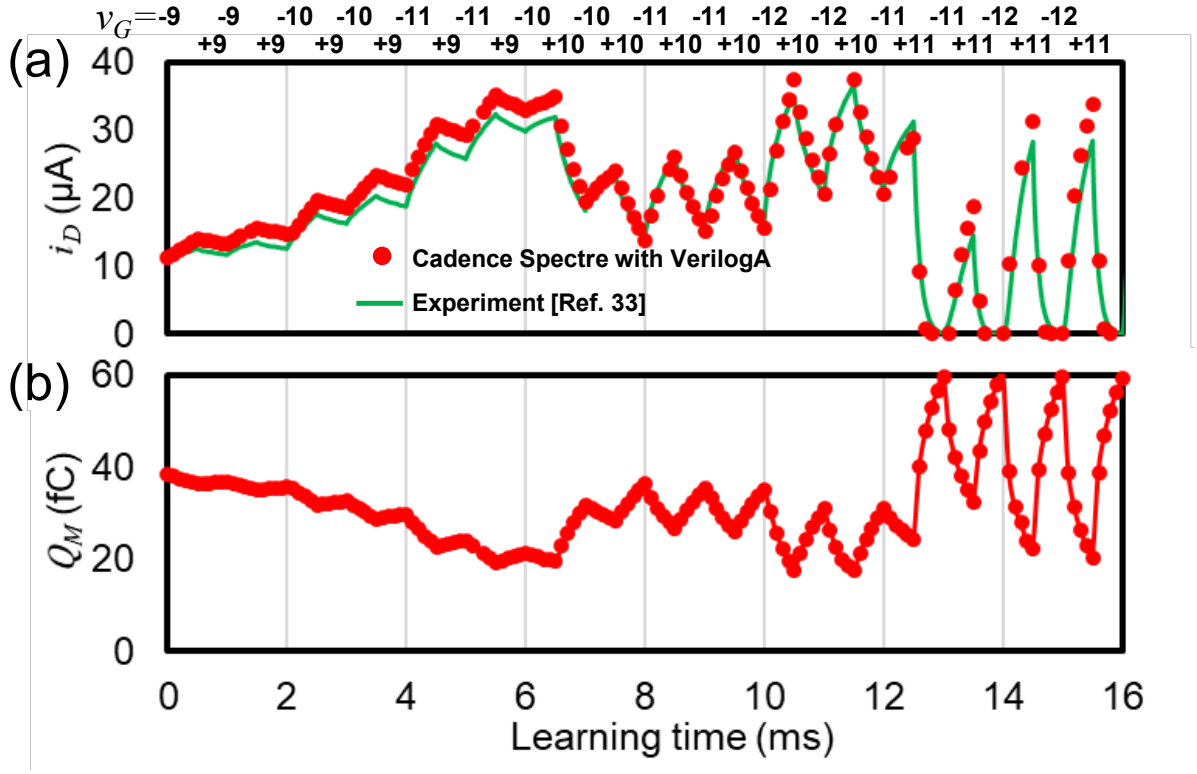

Figure 8. (a) Cadence Spectre simulation results for SONOS current, i_D, versus the accumulated learning time compared with the experimental results of Ref. [33] under v_G= 2.4 V and v_D= 0.1 V for reading the channel current. While the measurements of i_D for the experimental SONOS devices were conducted after every writing pulse (-9 ~ -12 V for potentiation, +9 ~ +11 V for depression with 10 μs width), our Cadence Spectre simulations were sampled after every 100 μs. Slight errors are attributed to the deviations from the perfectly exponential increase of M with Q_M assumed in our compact model. (b) The corresponding evolution of the state variable Q_M is available from the simulation, which demonstrates the strong correlation between ΔQ_M and Δi_D.

We deployed the compact model in a commercial circuit simulation package, Cadence Spectre, using VerilogA to assess the agreement with the measured drain current (with v_D=0.1 V and v_G=2.4 V during reading)[33] after various training pulses (v_G) of -12, -11, -10, -9, +9, +10, +11 V as shown in Fig. 8 (a). Despite the remarkable simplicity of the compact model, good agreement between the simulation and the experiment was confirmed with the accuracy higher than 90 %, proving that our compact model captures the essential physics in a simple way. Figure 8 (b) shows the corresponding evolution of the key state variable Q_M, where a strong correlation is observed between ΔQ_M and Δi_D as foreseen by the description of R_{CH} in Eq. (6).

In conclusion, this work presented a well-posed compact model of a SONOS three-terminal synaptic circuit element that can be readily utilized by circuit designers for neuromorphic computing circuits/systems. This compact model offers the rare combination of good predictability stemming from physics-driven state variables and simplicity. The basis of the compact model is rooted on the carrier concentration modulation in the conducting channel, so that it is readily applicable to other three-terminal synaptic circuit elements besides SONOS devices.

Acknowledgements

AAT and MJM were support by the Laboratory Directed Research and Development program at Sandia National Laboratories, a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525, and by the DOE Office of Science Research Program for Microelectronics Codesign (sponsored by ASCR, BES, HEP, NP, and FES) through the Abisko Project, PM Robinson Pino (ASCR).

References

1. LeCun, Y.; Bengio, Y.; Hinton, G., Deep learning. *Nature* **2015,** *521* (7553), 436-444.
2. Krizhevsky, A.; Sutskever, I.; Hinton, G., Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems* **2012,** *25*, 1097-1105.
3. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P., Gradient-based learning applied to document recognition. *J Proceedings of the IEEE* **1998,** *86* (11), 2278-2324.
4. LeCun, Y.; Boser, B.; Denker, J. S.; Henderson, D.; Howard, R. E.; Hubbard, W.; Jackel, L., Backpropagation applied to handwritten zip code recognition. *J Neural computation* **1989,** *1* (4), 541-551.
5. Shafiee, A.; Nag, A.; Muralimanohar, N.; Balasubramonian, R.; Strachan, J. P.; Hu, M.; Williams, R. S.; Srikumar, V., ISAAC: A convolutional neural network accelerator with in-situ analog arithmetic in crossbars. *J ACM SIGARCH Computer Architecture News* **2016,** *44* (3), 14-26.
6. Chen, Y.-H.; Krishna, T.; Emer, J. S.; Sze, V., Eyeriss: An energy-efficient reconfigurable accelerator for deep convolutional neural networks. *J IEEE journal of solid-state circuits* **2016,** *52* (1), 127-138.
7. Silver, D.; Schrittwieser, J.; Simonyan, K.; Antonoglou, I.; Huang, A.; Guez, A.; Hubert, T.; Baker, L.; Lai, M.; Bolton, A., Mastering the game of go without human knowledge. *Nature* **2017,** *550* (7676), 354-359.
8. Jouppi, N. P.; Young, C.; Patil, N.; Patterson, D.; Agrawal, G.; Bajwa, R.; Bates, S.; Bhatia, S.; Boden, N.; Borchers, A. In *In-datacenter performance analysis of a tensor processing unit*, Proceedings of the 44th annual international symposium on computer architecture, 2017; pp 1-12.
9. Williams, R. S., What's Next?[The end of Moore's law]. *J Computing in Science Engineering* **2017,** *19* (2), 7-13.
10. Strukov, D. B.; Snider, G. S.; Stewart, D. R.; Williams, R. S., The missing memristor found. *Nature* **2008,** *453* (7191), 80-83.
11. Chua, L., Memristor-the missing circuit element. *J IEEE Transactions on circuit theory* **1971,** *18* (5), 507-519.
12. Chua, L. O.; Kang, S. M., Memristive devices and systems. *J Proceedings of the IEEE* **1976,** *64* (2), 209-223.
13. Yang, J. J.; Zhang, M.-X.; Strachan, J. P.; Miao, F.; Pickett, M. D.; Kelley, R. D.; Medeiros-Ribeiro, G.; Williams, R. S., High switching endurance in TaO x memristive devices. *Applied Physics Letters* **2010,** *97* (23), 232102.
14. Kim, K. M.; Williams, R. S., A family of stateful memristor gates for complete cascading logic. *IEEE Transactions on Circuits Systems I: Regular Papers* **2019,** *66* (11), 4348-4355.
15. Prezioso, M.; Merrikh-Bayat, F.; Hoskins, B.; Adam, G. C.; Likharev, K. K.; Strukov, D., Training and operation of an integrated neuromorphic network based on metal-oxide memristors. *Nature* **2015,** *521* (7550), 61-64.
16. Kim, S.; Ishii, M.; Lewis, S.; Perri, T.; BrightSky, M.; Kim, W.; Jordan, R.; Burr, G.; Sosa, N.; Ray, A. In *NVM neuromorphic core with 64k-cell (256-by-256) phase change memory synaptic array with on-chip neuron circuits for continuous in-situ learning*, 2015 IEEE international electron devices meeting (IEDM), IEEE: 2015; pp 17.1. 1-17.1. 4.
17. Ambrogio, S.; Narayanan, P.; Tsai, H.; Shelby, R. M.; Boybat, I.; Di Nolfo, C.; Sidler, S.; Giordano, M.; Bodini, M.; Farinha, N., Equivalent-accuracy accelerated neural-network training using analogue memory. *Nature* **2018,** *558* (7708), 60-67.
18. Yi, S.-i.; Kumar, S.; Williams, R. S., Improved Hopfield Network Optimization using Manufacturable Three-terminal Electronic Synapses. *IEEE Transactions on Circuits and Systems I: Regular Papers* **2021,** *68* (12), 4970-4978.

19. Yi, S.-i.; Kim, J., Novel Program Scheme of Vertical NAND Flash Memory for Reduction of Z-Interference. *Micromachines* **2021,** *12* (5), 584.
20. Kim, H.; Park, J.; Kwon, M.-W.; Lee, J.-H.; Park, B.-G., Silicon-based floating-body synaptic transistor with frequency-dependent short-and long-term memories. *IEEE Electron Device Letters* **2016,** *37* (3), 249-252.
21. Mahmoodi, M. R.; Strukov, D. In *An ultra-low energy internally analog, externally digital vector-matrix multiplier based on NOR flash memory technology*, Proceedings of the 55th Annual Design Automation Conference, 2018; pp 1-6.
22. Hu, M.; Graves, C. E.; Li, C.; Li, Y.; Ge, N.; Montgomery, E.; Davila, N.; Jiang, H.; Williams, R. S.; Yang, J. J., Memristor-based analog computation and neural network classification with a dot product engine. *Advanced Materials* **2018,** *30* (9), 1705914.
23. Wang, Z.; Li, C.; Song, W.; Rao, M.; Belkin, D.; Li, Y.; Yan, P.; Jiang, H.; Lin, P.; Hu, M., Reinforcement learning with analogue memristor arrays. *Nature Electronics* **2019,** *2* (3), 115-124.
24. Li, C.; Wang, Z.; Rao, M.; Belkin, D.; Song, W.; Jiang, H.; Yan, P.; Li, Y.; Lin, P.; Hu, M., Long short-term memory networks in memristor crossbar arrays. *Nature Machine Intelligence* **2019,** *1* (1), 49-57.
25. Li, C.; Belkin, D.; Li, Y.; Yan, P.; Hu, M.; Ge, N.; Jiang, H.; Montgomery, E.; Lin, P.; Wang, Z., Efficient and self-adaptive in-situ learning in multilayer memristor neural networks. *Nature communications* **2018,** *9* (1), 1-8.
26. Wang, Z.; Joshi, S.; Savel'ev, S.; Song, W.; Midya, R.; Li, Y.; Rao, M.; Yan, P.; Asapu, S.; Zhuo, Y., Fully memristive neural networks for pattern classification with unsupervised learning. *Nature Electronics* **2018,** *1* (2), 137-145.
27. Danial, L.; Gupta, V.; Pikhay, E.; Roizin, Y.; Kvatinsky, S. In *Modeling a floating-gate memristive device for computer aided design of neuromorphic computing*, 2020 Design, Automation & Test in Europe Conference & Exhibition (DATE), IEEE: 2020; pp 472-477.
28. Zhang, C.; Hasan, S. R., A new floating-gate MOSFET model for analog circuit simulation and design. *Analog Integrated Circuits and Signal Processing* **2019,** *101* (1), 1-11.
29. Kim, M.; Kim, S.; Shin, H., A compact model for ISPP of 3-D charge-trap NAND flash memories. *IEEE Transactions on Electron Devices* **2020,** *67* (8), 3095-3101.
30. Fuller, E. J.; Keene, S. T.; Melianas, A.; Wang, Z.; Agarwal, S.; Li, Y.; Tuchman, Y.; James, C. D.; Marinella, M. J.; Yang, J. J., Parallel programming of an ionic floating-gate memory array for scalable neuromorphic computing. *Science* **2019,** *364* (6440), 570-574.
31. van de Burgt, Y.; Lubberman, E.; Fuller, E. J.; Keene, S. T.; Faria, G. C.; Agarwal, S.; Marinella, M. J.; Talin, A. A.; Salleo, A., A non-volatile organic electrochemical device as a low-voltage artificial synapse for neuromorphic computing. *Nature Materials* **2017,** *16* (4), 414-418.
32. Fuller, E. J.; Gabaly, F. E.; Léonard, F.; Agarwal, S.; Plimpton, S. J.; Jacobs-Gedrim, R. B.; James, C. D.; Marinella, M. J.; Talin, A. A., Li-ion synaptic transistor for low power analog computing. *Advanced Materials* **2017,** *29* (4), 1604310.
33. Agarwal, S.; Garland, D.; Niroula, J.; Jacobs-Gedrim, R. B.; Hsia, A.; Van Heukelom, M. S.; Fuller, E.; Draper, B.; Marinella, M., Using floating-gate memory to train ideal accuracy neural networks. *IEEE Journal on Exploratory Solid-State Computational Devices Circuits* **2019,** *5* (1), 52-57.
34. Widrow, B., *Adaptive" adaline" Neuron Using Chemical" memistors."*. 1960.
35. Adhikari, S. P.; Kim, H., Why are memristor and memistor different devices? In *Memristor networks*, Springer: 2014; pp 95-112.
36. Kim, H.; Adhikari, S. P., Memistor is not memristor [express letters]. *IEEE Circuits and Systems Magazine* **2012,** *12* (1), 75-78.

37. Jang, J.; Kim, H.-S.; Cho, W.; Cho, H.; Kim, J.; Shim, S. I.; Jeong, J.-H.; Son, B.-K.; Kim, D. W.; Shim, J.-J. In *Vertical cell array using TCAT (Terabit Cell Array Transistor) technology for ultra high density NAND flash memory*, 2009 Symposium on VLSI Technology, IEEE: 2009; pp 192-193.
38. Nowak, E.; Kim, J.-H.; Kwon, H.; Kim, Y.-G.; Sim, J. S.; Lim, S.-H.; Kim, D. S.; Lee, K.-H.; Park, Y.-K.; Choi, J.-H. In *Intrinsic fluctuations in vertical NAND flash memories*, 2012 Symposium on VLSI Technology (VLSIT), IEEE: 2012; pp 21-22.
39. Kang, D.; Kim, M.; Jeon, S. C.; Jung, W.; Park, J.; Choo, G.; Shim, D.-k.; Kavala, A.; Kim, S.-B.; Kang, K.-M. In *13.4 A 512Gb 3-bit/Cell 3D 6 th-Generation V-NAND flash memory with 82MB/s write throughput and 1.2 Gb/s interface*, 2019 IEEE International Solid-State Circuits Conference-(ISSCC), IEEE: 2019; pp 216-218.
40. Lee, S.; Kim, C.; Kim, M.; Joe, S.-m.; Jang, J.; Kim, S.; Lee, K.; Kim, J.; Park, J.; Lee, H.-J. In *A 1Tb 4b/cell 64-stacked-WL 3D NAND flash memory with 12MB/s program throughput*, 2018 IEEE International Solid-State Circuits Conference-(ISSCC), IEEE: 2018; pp 340-342.
41. Chua, L., Everything you wish to know about memristors but are afraid to ask. In *Handbook of Memristor Networks*, Springer: 2019; pp 89-157.
42. Sedra, A. S.; Smith, K. C.; Carusone, T. C.; Gaudet, V., *Microelectronic circuits*. Oxford university press New York: 2004; Vol. 4.
43. Streetman, B. G.; Banerjee, S., *Solid state electronic devices*. Pearson/Prentice Hall Upper Saddle River: 2006; Vol. 10.
44. Synposys, "Sentaurus Manual S-Device" L-version. **2016**.
45. Ascoli, A.; Tetzlaff, R.; Chua, L. O.; Strachan, J. P.; Williams, R. S. J. I. T. o. C.; Papers, S. I. R., History erase effect in a non-volatile memristor. **2016,** *63* (3), 389-400.
46. Pickett, M. D.; Strukov, D. B.; Borghetti, J. L.; Yang, J. J.; Snider, G. S.; Stewart, D. R.; Williams, R. S., Switching dynamics in titanium dioxide memristive devices. *Journal of Applied Physics* **2009,** *106* (7), 074508.
47. Danesh, C. D.; Shaffer, C. M.; Nathan, D.; Shenoy, R.; Tudor, A.; Tadayon, M.; Lin, Y.; Chen, Y., Synaptic resistors for concurrent inference and learning with high energy efficiency. *Advanced Materials* **2019,** *31* (18), 1808032.
48. Li, Y.; Fuller, E. J.; Sugar, J. D.; Yoo, S.; Ashby, D. S.; Bennett, C. H.; Horton, R. D.; Bartsch, M. S.; Marinella, M. J.; Lu, W., Filament-Free Bulk Resistive Memory Enables Deterministic Analogue Switching. *Advanced Materials* **2020,** *32* (45), 2003984.
49. Compagnoni, C. M.; Goda, A.; Spinelli, A. S.; Feeley, P.; Lacaita, A. L.; Visconti, A., Reviewing the evolution of the NAND flash technology. *Proceedings of the IEEE* **2017,** *105* (9), 1609-1633.

WILEY-VCH

Supplementary Information

Physical Compact Model for Three-Terminal SONOS Synaptic Circuit Element

Su-in Yi[1], Alec Talin,[2] Matthew Marinella[3], and R. Stanley Williams[1,*]

[1]Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA

[2]Sandia National Laboratories, Livermore, CA 94550, USA

[3]Arizona State University, Tempe, AZ 85281, USA

* Corresponding author: rstanleywilliams@tamu.edu

1. Reduction of R_{CH} under three timescales

The full description of R_{CH} in our compact model, which relies on both time constants, RC and MC_M, can be written as

$$R_{CH} = \left[\mu \frac{1}{2L^2}\left(C(v_G - V_T - v_M)\left(1 - e^{-t/(RC)}\right) - \chi C_M (v_G - V_T - v_M)\left(1 - e^{-t/(MC_M)}\right)\right)\right]^{-1} \quad \text{(S1)}$$

As an expansion of the quasi-static expression for the drain current of a MOSFET as discussed with Eq. (3), Eq. (S1) can be specialized to three time intervals

$$R_{CH} = \left[\mu \frac{1}{2L^2}\left(C(v_G - V_T - v_M)\left(1 - e^{-t/(RC)}\right)\right)\right]^{-1} \quad \text{when} \quad t < 1 \text{ ns} \quad \text{(S2)}$$

$$R_{CH} = \left[\mu \frac{1}{2L^2}\left(C(v_G - V_T - v_M)\right)\right]^{-1} \quad \text{when } 1 \text{ ns} < t < 1 \text{ µs}^{1,2} \quad \text{(S3)}$$

$$R_{CH} = \left[\mu \frac{1}{2L^2}\left(C(v_G - V_T - v_M) - \chi C_M (v_G - V_T - v_M)\left(1 - e^{-t/(MC_M)}\right)\right)\right]^{-1} \quad \text{when } 1 \text{ µs} < t \quad \text{(S4)}$$

due to the significant contrast (~ three orders of magnitude) between RC and MC_M. Therefore, for practical circuit simulations as a synaptic circuit element (under training), one may employ Eq. (S4), where Q becomes a parameter rather than a state variable, so that the computing time for simulations can be minimized[3]. It also resolves the issue of choosing a large resistance such as R=10^8 Ω to manage the simulation time (timestep as large as 1 ns) in our work.

2. Derivation of relationship between ln(Q_{max}-Q) and t for an R-C circuit

The Dynamic Route Map (DRM) as shown in Fig. S2 of a series R-C circuit under a voltage source is

$$\frac{dQ}{dt} = -\frac{Q}{RC} + \frac{V}{R} \tag{S5}$$

for which the general solution is

$$Q = A \cdot \exp\left(-\frac{t}{RC}\right) + CV \ , \tag{S6}$$

where A depends on the initial condition, Q at t=0. Eq. (S6) is equivalent to

$$-\frac{t}{RC} = \ln\left(\frac{Q-CV}{A}\right) \ , \tag{S7}$$

which can be rearranged, depending on the sign of V, as

$$-\frac{t}{RC} = \begin{cases} \ln(CV - Q) - \ln(-A) & if\ V > 0 \\ \ln(Q - CV) - \ln(A) & if\ V < 0 \end{cases} \tag{S8}$$

where CV is equal to the maximum charge, Q_{max}, under a bias, V.

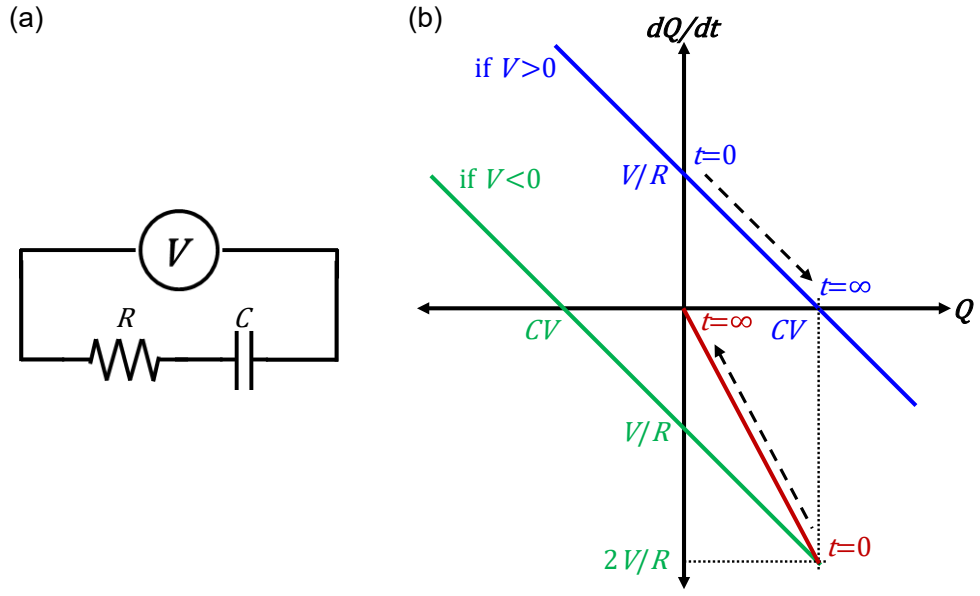

Figure S1. (a) A series R-C circuit under a DC voltage bias, (b) DRM of an R-C circuit when V>0 (blue), V<0 (green), and V<0 with a unipolar capacitor (red) similar to C_M of a SONOS.

3. Extraction of the long-term gate current ($I_{G,0}$) from the mixture ($I_{G,0}+I_{G,\infty}$)

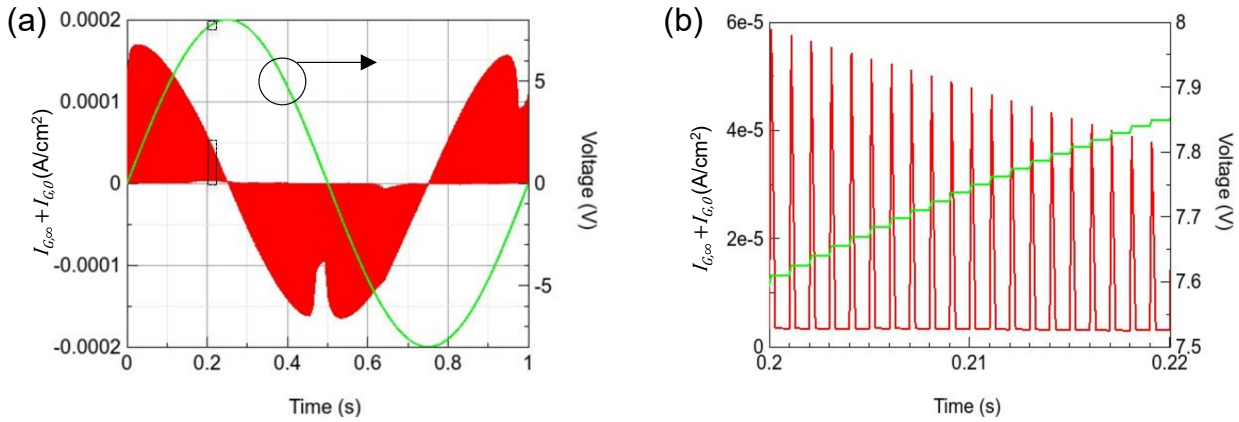

Figure S2. (a) Total gate current (the sum of the long-term, $I_{G,0}$, and the short-term dynamics, $I_{G,\infty}$ with an incremental AC voltage obtained from a TCAD simulation. (b) Magnified view of (c) for $0.2<t<0.22$, showing the incremental increase of voltage with a ramp-up time of 0.1 ms followed by a holding time of 0.9 ms to separate $I_{G,0}$ and $I_{G,\infty}$.

1. Sedra, A. S.; Smith, K. C.; Carusone, T. C.; Gaudet, V., *Microelectronic circuits*. Oxford university press New York: 2004; Vol. 4.
2. Streetman, B. G.; Banerjee, S. K., Solid State Electronic Devices: Global Edition. Pearson education: 2016.
3. Yi, S.-i.; Kumar, S.; Williams, R. S., Improved Hopfield Network Optimization using Manufacturable Three-terminal Electronic Synapses. *IEEE Transactions on Circuits and Systems I: Regular Papers* **2021,** *68* (12), 4970-4978.